Act report

Project Description:

In this project, I was responsible to extract some data from twitter API, the data was some reviews from **@dog_rates** who known also as **WeRateDogs**, and gather data from the tweet archive of Twitter, and get the data to analyze it later using an image regression model to predict the dog's breed.

The data I get was stored in two CSV files, the first one has data about the tweets (tweet ID, timestamp, text, etc.) the data was about 2356 row, and another file that has data about the dogs itself (rating, breed, name).

This data has 4 main dog's breeds, they are (Pupper, Doggo, Puppo, Floofer), and the chart for this breed is in figure (1) and (2):
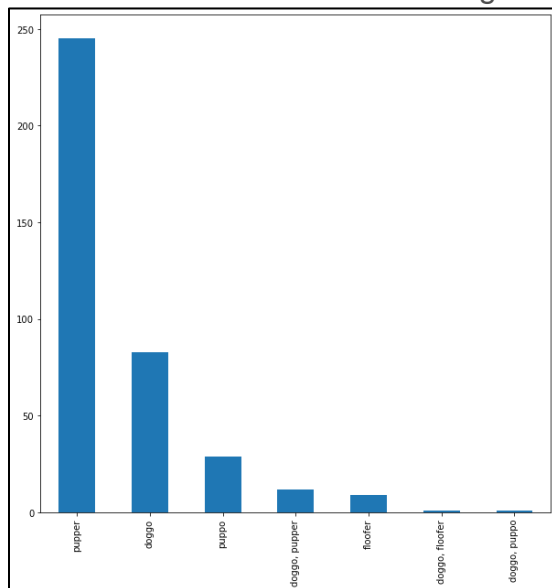

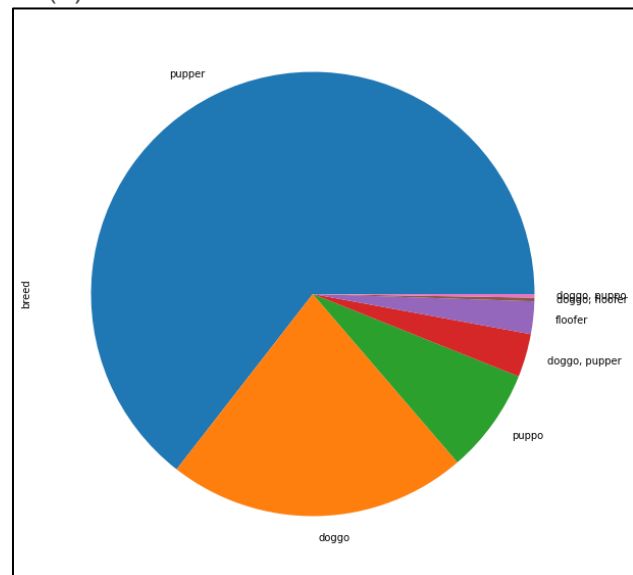*Figure 1: Bar chart to show the number of dogs in each breed*

Figure 2: pie chart to show the number of dogs in each breed

Project Data

As I mentioned above you need to gather tweeter API to get the data, one of very useful tool enabling you to do this mission is **Tweepy** library, this library helps you to get data from tweeter, you need first to have email on **Twitter's Developer Portal**, and have your API key to get your data in json file, then you can extract your data from it and do your wrangling process on it to get finally a clean data enable you to add to any ML or DL model and in the next figures is the screenshot of the first rows in my data.

	rating_numerator	rating_denominator	name	breed
0	13	10	Phineas	NaN
1	13	10	Tilly	NaN
2	12	10	Archie	NaN
3	13	10	Darla	NaN
4	12	10	Franklin	NaN
5	13	10	NaN	NaN
6	13	10	Jax	NaN
7	13	10	NaN	NaN
8	13	10	Zoey	NaN
9	14	10	Cassie	doggo

Figure 3: The first 10 rows in dog's information

	tweet_id	in_reply_to_status_id	in_reply_to_user_id	timestamp	source	text
0	892420643555336193	NaN	NaN	2017-08-01 16:23:56 +0000	<a href="http://vine.com" rel="nofollow">Vine …	This is Phineas. He's a mystical boy. Only eve…
1	892177421306343426	NaN	NaN	2017-08-01 00:17:27 +0000	<a href="http://vine.com" rel="nofollow">Vine …	This is Tilly. She's just checking pup on you.…
2	891815181378084864	NaN	NaN	2017-07-31 00:18:03 +0000	<a href="http://vine.com" rel="nofollow">Vine …	This is Archie. He is a rare Norwegian Pouncin…
3	891689557279858688	NaN	NaN	2017-07-30 15:58:51 +0000	<a href="http://vine.com" rel="nofollow">Vine …	This is Darla. She commenced a snooze mid meal…
4	891327558926688256	NaN	NaN	2017-07-29 16:00:24 +0000	<a href="http://vine.com" rel="nofollow">Vine …	This is Franklin. He would like you to stop ca…

Figure 4: The first 5 rows in tweets information Part1

retweeted_status_id	retweeted_status_user_id	expanded_urls
NaN	NaN	https://twitter.com/dog_rates/status/892420643…
NaN	NaN	https://twitter.com/dog_rates/status/892177421…
NaN	NaN	https://twitter.com/dog_rates/status/891815181…
NaN	NaN	https://twitter.com/dog_rates/status/891689557…
NaN	NaN	https://twitter.com/dog_rates/status/891327558…

Figure 5: The first 5 rows in tweets information Part2